Journal of Public Economics 160 (2018) 33–49

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Be careful what you calibrate for: Social discounting in general equilibrium

Lint Barrage

Brown University, United States
NBER, United States

ARTICLE INFO

Keywords:
Discounting
Social discount rate
Carbon taxes
Ramsey taxation
Distortionary taxes
General equilibrium climate-economy model

ABSTRACT

Concerns about intergenerational equity have led to an influential practice of setting social utility discount rates based on ethical considerations rather than to match household behavior, particularly in climate change economics (e.g., Stern, 2006). This paper formalizes the broader policy implications of this approach in general equilibrium by characterizing jointly optimal environmental and fiscal policies in a climate-economy model with differential planner-household discounting. First, I show that decentralizing the optimal allocation requires not only high carbon prices but also fundamental changes to tax policy: If the government discounts the future less than households, implementing the optimal allocation requires an effective capital income subsidy (a negative intertemporal wedge), and, in a setting with distortionary taxation, an effective labor-consumption tax wedge that is decreasing over time. Second, if the government cannot subsidize capital income, the *constrained-optimal* carbon tax may be up to 50% below the present value of marginal damages (the social cost of carbon) due to the general equilibrium effects of climate policy on household savings. Third, given the choice to optimize either carbon, capital, or labor income taxes, the socially discounting planner's welfare ranking is ambiguous over a standard range of parameters. Overall, in general equilibrium, a policy-maker's choice to adopt differential social discounting may thus overturn conventional recommendations for both environmental and fiscal policy.

1. Introduction

Economists have long debated the appropriate social rate of discount with which policy-makers should value the future.[1] Recently, in the context of the economics of climate change, the question of social discounting has again risen to the forefront of academic and policy debates (e.g., Arrow et al., 1996, 2012). As the impacts of greenhouse gas emissions occur over long time horizons, optimal climate policy depends critically on discounting parameters (Nordhaus, 2007, 2008; Interagency Working Group, 2010). One influential view posits that it is "ethically indefensible" to discount the utility of future generations, as famously stated by Ramsey (1928). Economists embracing this view have incorporated near-zero pure rates of social time preference into their models, yielding aggressive climate policy recommendations (Stern, 2006; Cline, 1992). This approach has been controversial for several reasons, including the fact that it does not align with households' intertemporal preferences as revealed through their savings behavior (Nordhaus, 2007; Dasgupta, 2008). Indeed, Ramsey himself noted that " the rate of saving which the [zero-discounting optimality] rule requires is greatly in excess of that which anyone would normally suggest " (Ramsey, 1928). While Ramsey's ethical critique is widely cited in support of near-zero discounting (e.g., Interagency Working Group, 2010), his work also reminds us that such discounting does not match household behavior. Several studies have thus proposed frameworks where agents and the planner discount utility differently (e.g., Farhi and Werning, 2005, 2007, 2010; Kaplow et al., 2010; Goulder and Williams, 2012; von Below, 2012; Belfiori, 2017). This approach can be formally microfounded by an overlapping generations model where the planner places a higher welfare weight on future generations above the current generation's private altruism towards the future (Bernheim, 1989; Farhi and Werning, 2005, 2007; Belfiori, 2017; described below). While social discounting can thus be used to address intergenerational ethical concerns in climate policy, it should also be expected to have implications for, e.g., capital allocations and fiscal policy. Though the potential for such broader implications has been frequently noted (e.g., Manne, 1995; Caplin and Leahy, 2004; Goulder and Williams, 2012), few studies have formalized these effects to date.

This paper thus formalizes and quantifies the broader policy implications of differential discounting in a dynamic general equilibrium climate-economy model. I theoretically characterize jointly optimal

E-mail address: lint_barrage@brown.edu.

[1] This vast literature ranges from debates on the fundamental principles of what discounting should capture (e.g., Baumol, 1968; Sen, 1982; Lind, 1982) to modern treatments about the implications of factors such as heterogeneity (e.g., Gollier and Zeckhauser, 2005; Heal and Millner, 2013), uncertainty (e.g., Newell and Pizer, 2003; Gollier and Weitzman, 2010), the structure of preferences (e.g., Caplin and Leahy, 2004), and commitment (Sleet and Yeltekin, 2006), inter alia.

https://doi.org/10.1016/j.jpubeco.2018.02.012
Received 26 June 2016; Received in revised form 11 November 2017; Accepted 23 February 2018
Available online 15 March 2018
0047-2727/ © 2018 Elsevier B.V. All rights reserved.

climate and fiscal policy across three settings: a first-best setting, a second-best setting where the planner is constrained in his ability to subsidize capital income, and a second-best fiscal setting in the Ramsey tradition where the planner must raise revenues from distortionary taxes. I then quantify optimal policies and welfare by integrating differential discounting into two climate-economy models: the generalized numerical implementation of Golosov, Hassler, Krusell, and Tsyvinski's model (2014, "GHKT") by Barrage (2014), and the COMET model (Barrage, 2016) which extends the seminal DICE framework of Nordhaus (see, e.g., Nordhaus, 2008, 2010a,b) to incorporate distortionary taxation and government expenditures.

The approach of this paper seeks to contribute to the prior literature in several ways. On the one hand, both von Below (2012) and Belfiori (2017) study differential discounting in general equilibrium climate-economy models, but focus on energy and climate policy implications in a first-best setting where the planner can freely subsidize capital income (von Below, 2012) or where oil is the only factor of production (Belfiori, 2017), so that there is no role for fiscally constrained or broader tax policy. On the other hand, Farhi and Werning (2005, 2010) study differential discounting and fiscal policy, but focus on a Mirrleesian economy (without climate change). In this setting, tax policy is designed to trade off insurance and incentives for agents that experience productivity shocks, and face the risk of being born into families with less productive parents. Distortions arise as the planner cannot observe households' work effort or productivity. In contrast, in the Ramsey setting considered here, distortions arise as the government must raise revenues but cannot impose lump-sum taxes. The main results of the analysis are as follows.

First, differential discounting fundamentally alters optimal tax policy prescriptions. In a first-best setting, decentralizing the optimal allocation requires both high carbon taxes and capital income subsidies. Intuitively, this is because households are too impatient from the planner's perspective, and consequently fail to invest sufficiently in the economy's capital stocks. While this insight was previously formalized by von Below (2012) in the first-best, I further show that it holds even in a fiscal setting where government revenues must be raised from distortionary taxes. In this case, if the government values the future more than households, decentralizing the optimal allocation requires not only capital income subsidies but also an effective labor-consumption tax wedge that is decreasing over time. These policies stand in stark contrast to the classic Ramsey prescriptions of zero capital income taxes (i.e., no intertemporal distortions), and constant (smooth) intratemporal distortions through, e.g., constant labor income taxes (see, e.g., Chari and Kehoe, 1999; Atkeson et al., 1999). These changes are moreover quantitatively significant: Adopting the pure rate of social time preference advocated by Stern (2006) (0.1% per year) yields optimal capital income subsidies ranging from 30% to 75%, and changes optimal labor income taxes from a constant rate of 41% with standard discounting to a high initial rate of 53% that decreases to 33% by the end of the century. Though different from the standard Ramsey setting, the results broadly align with Farhi and Werning's (2005, 2010) finding that – in a Mirrleesian economy with social discounting – the constrained-efficient allocation can be decentralized by negative marginal estate taxes, indicating that bequests should be subsidized. While Mirrleessian and Ramsey economies generally lead to different tax policy recommendations, with differential discounting, the results of this paper suggest that both frameworks broadly agree on the desirability of subsidies to increase the returns to savings.

In reality, most countries *tax* capital income, calling into question the political feasibility of capital income subsidies. Second, I thus consider policy design from the perspective of a socially discounting planner who cannot subsidize capital income. Given this basic constraint, I find that even a planner adopting the social discounting preferences advocated by Stern (2006) may not want to impose a 'Sternian' carbon tax (i.e., a tax equal to the socially discounted marginal damages of emissions) if he cannot subsidize savings at the same time. More formally, the constrained-optimal carbon tax must be adjusted for the general equilibrium effects of climate policy on households' incentives to save: (i) If climate change decreases the returns to capital, a component must be added to the optimal carbon tax, ceteris paribus. (ii) If energy and capital are complements in production, a component must be subtracted from the optimal carbon tax, ceteris paribus. Intuitively, without a capital income subsidy, private returns to savings are too low. To the extent that climate policy decreases (increases) the returns to capital, it therefore exacerbates (mitigates) this inefficiency, and must be adjusted accordingly.[2] Theoretically, the net sign of these adjustments is ambiguous. Quantitatively, I find that optimal carbon taxes are 5–40 % *lower* if the planner cannot subsidize capital income, and 10–60 % lower if the planner cannot reduce capital income taxes below a positive level of 30%. The largest downward adjustments occur when energy and capital are complements in production, as higher energy prices reduce the marginal product of capital. Since the planner cares about the overall level of assets given to future generations – both natural and man-made – he adjusts carbon taxes downward to mitigate their undesirable effects on capital accumulation. While the size of this adjustment is sensitive to the parameters, the results imply that a constrained fiscal environment may significantly decrease optimal carbon taxes even from the perspective of a socially discounting planner.

Finally, I compare the welfare gains that a planner with a near-zero pure rate of social time preference would achieve by optimizing either carbon, capital income, or labor income taxes. Perhaps surprisingly, the relative importance of these policy levers is ambiguous, and depends critically on the parameters. For example, the benchmark calibration assumes a consumption elasticity of $\sigma = 1.5$. In this setting, the welfare benefits from adjusting carbon taxes from standard to socially discounted levels (+0.96% permanent consumption increase equivalent) are lower than those from adjusting either capital income taxes (+1.36%) or labor income taxes (+2.30%). In contrast, a planner with logarithmic preferences would indeed achieve significantly larger welfare gains from adjusting climate policy than fiscal policy. While these calculations abstract from many important complexities and are subject to many caveats, they do illustrate that climate protection is only one of various investments society can make to benefit future generations (e.g., health, education, and private capital). A government that weighs the utility of future households more highly than its citizens should thus incentivize larger investments in an appropriately balanced portfolio of all such assets.

Of course there are a number of important reasons why investments in the climate may still need to be discounted differently from other assets in such a portfolio. For example, climate investments accrue benefits over much longer time horizons than other assets, and recent empirical evidence points to a significantly downward-sloping term structure of discount rates over such time horizons (Giglio et al., 2014). Several underlying factors could generate this pattern, including uncertainty over growth (see, e.g., Gollier, 2016, Gollier and Weitzman, 2010) or the structure of household preferences. The results of this paper highlight the potential importance of formally incorporating such microfoundations into macro-climate-economy models in order to assess their policy implications comprehensively.

The remainder of this paper proceeds as follows. Section 2 sets up the benchmark model and describes policies that decentralize the optimal allocation in the first-best. Section 3 characterizes constrained-optimal policy for a planner who cannot subsidize capital income. Section 4 presents the analysis in a setting with distortionary taxes. Section 5 summarizes the calibration and the quantitative results.

[2] One could also ask how capital should be taxed when carbon cannot be priced. For a treatment of this issue in a decentralized economy with altruistic households that invest in dirty capital to protect their offspring against climate change (but thereby exacerbate the externality), see Asheim and Nesje (2016).

Finally, Section 6 concludes.

2. Benchmark theoretical model

2.1. Social discounting setup

Following recent literature (e.g., Farhi and Werning, 2005, 2007; Belfiori, 2017), I model social discounting as a disagreement in utility discount factors between a dynastic household (β_H) and the social planner (β_G). As shown by these previous studies and Bernheim (1989), this specification can be microfounded by an overlapping generations framework where households live for one period and attach altruistic weight β_H to their children's welfare, but where the social planner *additionally* places a direct welfare weight α_t on each future generation t. This subsection reproduces their argument formally to showcase the intuition in this setting. First, from the perspective of private households alive at time t, their welfare v_t is given by:

$$v_t = U(C_t, S_t) + \beta_H v_{t+1} \tag{1}$$

where C_t is the aggregate consumption good and S_t represents the stock of carbon concentrations at time t. Iterating Eq. (1) forward yields the standard dynastic representation of household utility:

$$v_t = \sum_{s=0}^{\infty} (\beta_H)^s U(C_{t+s}, S_{t+s}). \tag{2}$$

Assume now that, in addition to indirectly valuing future generations through the current generation's altruism, the social planner also places a direct welfare weight α_t on each generation t, so that the social welfare function is given by:

$$W_0 \equiv \sum_{t=0}^{\infty} \alpha_t v_t = \sum_{t=0}^{\infty} \alpha_t \sum_{s=t}^{\infty} (\beta_H)^{s-t} U(C_s, S_s) \tag{3}$$

$$= \sum_{s=0}^{\infty} \left(\sum_{t=0}^{s} \alpha_t (\beta_H)^{s-t} \right) U(C_s, S_s). \tag{4}$$

Bernheim (1989) proves that, for appropriately set welfare weights, Eq. (4) corresponds to a time-consistent social planner's problem maximizing dynastic utility with social utility discount factor exceeding β_H. His results specifically imply that, for $\beta_G > \beta_H$,

$$\left(\sum_{t=0}^{s} \alpha_t (\beta_H)^{s-t} \right) = \alpha_0 (\beta_G)^s. \tag{5}$$

The weights $\alpha_t = (\beta_G)^t(\beta_G - \beta_H)\alpha_0$ are set decline geometrically in order to ensure dynamic consistency in the planner's problem, and the initial weight scalar $\alpha_0 = \frac{1-\beta_G}{1-\beta_H}$ is set to ensure that $\sum_{t=0}^{\infty} \alpha_t = 1$.[3] Consequently, if the social planner places positive weight on future generations' welfare above and beyond the current living population's private altruism, the corresponding social welfare function Eq. (3) can be represented by:

$$W_0 = \alpha_0 \left\{ \sum_{s=0}^{\infty} (\beta_G)^s U(C_s, S_s) \right\}. \tag{6}$$

Following prior studies (Farhi and Werning, 2005, 2007; Belfiori, 2017), I thus model the social discounting problem as one where the dynastic household maximizes Eq. (2) but the social planner seeks to maximizes Eq. (6).

[3] For detailed treatments of optimal climate policy with time-inconsistent preferences and limited commitment see, e.g., Gerlagh and Liski (2018) and Iverson and Karp (2017).

2.2. Decentralized economy

2.2.1. Households

Households earns income from labor L_t (inelastically supplied in the benchmark but endogenized in Section 4), return r_{t+1} on capital holdings K_t (net of depreciation δ and capital income tax τ_{kt}), and profits Π_t from the energy production sector, with associated flow budget constraint:

$$C_t + K_{t+1} \leq w_t L_t + \{1 + (r_t - \delta)(1 - \tau_{kt})\}K_t + \Pi_t + \Lambda_t \tag{7}$$

where Λ_t denotes government lump-sum transfers. This benchmark version of the model abstracts from consumption and labor taxes as they are not needed to decentralize the optimal allocation. That is, as the only distortions in the economy are the climate externality and the difference in planner and household discount rates, the two tax instruments considered - capital and carbon taxes - are sufficient to decentralize the optimal allocation. The household seeks to maximize Eq. (2) subject to Eq. (7), yielding the following savings optimality condition:

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H \{1 + (r_{t+1} - \delta)(1 - \tau_{kt+1})\} \tag{8}$$

where U_{ct} denotes the marginal utility of consumption in period t.

2.2.2. Production

Aggregate output Y_t is produced competitively using capital K_t, labor L_t^Y, and energy E_t inputs in a constant returns to scale (CRS) technology that satisfies the usual Inada conditions. Output further depends on carbon concentrations S_t and technology A_t:

$$Y_t = F(A_t(S_t); K_t, L_t^Y, E_t). \tag{9}$$

Profit-maximizing firms equate marginal products with factor prices in equilibrium:

$$\begin{aligned} F_{Lt} &= w_t \\ F_{Kt} &= r_t \\ F_{Et} &= p_{Et} \end{aligned} \tag{10}$$

where F_{jt} denotes the marginal product of input j at time t, and p_{Et} is the price of energy. Final output can be consumed or invested, with associated aggregate resource constraint:

$$C_t + K_{t+1} = Y_t + (1 - \delta)K_t. \tag{11}$$

The energy input E_t is assumed to be produced with labor inputs L_t^E and CRS technology:[4]

$$E_t = G_t(L_t^E). \tag{12}$$

Labor is perfectly mobile across sectors ($L_t = L_t^Y + L_t^E$), implying that wages will be equated in equilibrium. Finally, the statutory incidence of excise carbon fuel taxes τ_{Et} is on energy producers, whose profit maximization problem and condition in each period t are given by:

$$\begin{aligned} &\max_{L_t^E} (p_{Et} - \tau_{Et}) G_t(L_t^E) - w_t L_t^E \\ &\Rightarrow (p_{Et} - \tau_{Et}) G_{Lt} = w_t. \end{aligned} \tag{13}$$

2.2.3. Environment

Atmospheric carbon concentrations S_t are a function $\widetilde{S_t}$ of initial concentrations S_0 and emissions E_t dating back to the start of industrialization at time $-T$:

$$S_t = \widetilde{S_t}(S_0, E_{-T}, E_{-T+1}, ..., E_t) \tag{14}$$

with $\frac{\partial \widetilde{S_{t+j}}}{\partial E_t} \geq 0 \; \forall \; j \geq 0$. Note that energy inputs E_t are measured in tons

[4] See Belfiori (2017) for a detailed treatment of non-renewable resource dynamics with differential discounting.

of carbon-equivalent.

2.3. Planner's problem: first-best policy

In the first-best setting the planner has access to an unrestricted set of policy instruments, and maximizes socially discounted utility (Eq. (6)) subject to aggregate production (Eq. (9)), the resource constraint (Eq. (11)), the climate change constraint (Eq. (14)), energy production (Eq. (12)), and the labor resource constraint. The optimal allocation can then be characterized as follows. First, combine the planner's first-order conditions (FOCs) with respect to energy E_t, carbon concentrations S_t, and consumption C_t to obtain the standard optimality condition that the social marginal costs and benefits of carbon energy usage be equated:

$$\underbrace{F_{Et}}_{\substack{\text{Marginal}\\ \text{product of}\\ \text{energy}}} - \underbrace{\sum_{j=0}^{\infty} \beta_G^j \frac{\partial S_{t+j}}{\partial E_t}\left[\frac{\partial Y_{t+j}}{\partial S_{t+j}}\frac{U_{ct+j}}{U_{ct}} + \frac{U_{S_{t+j}}}{U_{ct}}\right]}_{\substack{\text{Marginal}\\ \text{damages from}\\ \text{carbon emissions}}} = \underbrace{\frac{w_t}{G_{Lt}}}_{\substack{\text{MC of}\\ \text{energy production}\\ (\text{in} C_t \text{ units})}}. \tag{15}$$

Next, combining the planner's FOCs with respect to consumption C_t and capital K_{t+1} yields the planner's Euler equation:

$$[F_{Kt+1} + (1-\delta)] = \frac{1}{\beta_G}\frac{U_{ct}}{U_{ct+1}}. \tag{16}$$

Comparing the planner's optimality conditions with those governing the behavior of households (Eq. (8)) and energy producers (Eq. (13)) in decentralized equilibrium, it is easy to show the following:

Result 1. The first-best allocation can be decentralized by the combination of a carbon tax set equal to the socially discounted value (SDV) of marginal damages of emissions at the optimum,

$$\tau_{Et}^* = \sum_{j=0}^{\infty} \beta_G^j \frac{\partial S_{t+j}}{\partial E_t}\left[\frac{U_{ct+j}}{U_{ct}}\frac{-\partial Y_{t+j}}{\partial S_{t+j}} + \frac{-U_{S_{t+j}}}{U_{ct}}\right] \tag{17}$$

and a capital income subsidy to equate the social and private marginal returns on savings:

$$\tau_{kt+1}^* = \left(\frac{\beta_H - \beta_G}{\beta_H}\right)\left(\frac{F_{Kt+1} + 1 - \delta}{F_{Kt+1} - \delta}\right). \tag{18}$$

Proof. See Appendix A. With standard discounting where the government adopts household preferences ($\beta_G = \beta_H$), the optimal capital income tax is zero. If, however, the planner weighs the future more highly than households, $\beta_G > \beta_H$, then the optimal capital income tax is negative, $\tau_{kt+1}^* < 0$. This theoretical need for a capital income subsidy with differential discounting was previously formalized by von Below (2012) and also mentioned by e.g., GHKT. However, this prescription stands in stark contrast to the common policy practice of taxing capital income. The next two sections thus introduce two layers of realism into the analysis: (1) constraints on the government's ability to subsidize capital income, and (2) government revenue requirements that must be met through distortionary taxes due to constraints on the government's ability to impose lump-sum taxes. Both model features lead to fundamentally different environmental and tax policy prescriptions with social discounting compared to classic results in the literature.

3. Constrained-optimal climate policy

The previous section showed that, if the government discounts the future less than households, then achieving the first-best allocation theoretically requires capital income subsidies. In reality, however, most countries tax capital income, and both academic debates and policy decisions surrounding carbon pricing are often made separately from fiscal policy. Consequently, this section characterizes *constrained-optimal* climate policy from the perspective of a planner who believes in social discounting but cannot subsidize savings. Since the social planner's problem is set up in primal form, the corresponding capital income tax constraint needs to be expressed in terms of allocations as well (see, e.g., Chari and Kehoe, 1999). In order to derive the constraint, first note that, in equilibrium (where $r_{t+1} = F_{Kt+1}$ as per Eq. (10)), the household's intertemporal optimality condition is given by:

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H\{1 + (F_{Kt+1} - \delta)(1 - \tau_{kt+1})\}. \tag{19}$$

The constraint that capital income cannot be subsidized ($\tau_{kt+1} \geq 0$) implies that:

$$\{1 + (F_{Kt+1} - \delta)(1 - \tau_{kt+1})\} \leq \{1 + (F_{Kt+1} - \delta)\}. \tag{20}$$

Given households' equilibrium behavior (Eq. (19)), the no-subsidy constraint Eq. (20) thus maps into:

$$\frac{U_{ct}}{\beta_H U_{ct+1}} \leq \{1 + (F_{Kt+1} - \delta)\}. \tag{21}$$

At this stage, Eq. (21) is the only constraint added to the planner's problem; other taxes – including lump-sum transfers – remain unrestricted. This setup ensures that any deviation of the optimal carbon tax from the Pigouvian levy Eq. (17) is driven solely by the unavailability of capital income subsidies in a setting with social discounting. In contrast, with an additional restriction that the planner cannot impose lump-sum taxes, it is well-known that the optimal carbon tax would differ from the Pigouvian levy even in a setting with standard discounting (see, e.g., Bovenberg and Goulder, 1996) where Eq. (21) would not be binding. The Online Appendix nonetheless shows that the constrained-optimal carbon tax formulation derived below is structurally robust to consideration of a Ramsey second-best fiscal setting (i.e., to the addition of constraints on lump-sum taxation) for $t > 0$. The general fiscal policy implications of social discounting in a second-best Ramsey setting are analyzed separately in Section 4.

In addition to Eq. (21), I now assume that carbon and consumption enter the household's utility separably. The Online Appendix formally shows that non-separability does not change the results presented below, but may add another term to the optimal carbon tax formulation.[5] Given these assumptions, Proposition 1 summarizes the results.

Proposition 1. *The constrained-optimal carbon tax in period $t > 0$ is implicitly defined by the sum of three parts: (i) the SDV of marginal climate damages to utility and production, (ii) the SDV of marginal climate damages to the marginal product of capital, and (iii) the social cost of marginal energy input use changes on the marginal product of capital:*

$$\tau_{Et}^* = \underbrace{\sum_{j=0}^{\infty} \beta_G^j \frac{\partial S_{t+j}}{\partial E_t}\left[\frac{-\partial Y_{t+j}}{\partial S_{t+j}}\frac{\lambda_{t+j}}{\lambda_t} + \frac{-U_{St+j}}{\lambda_t}\right]}_{\substack{\text{Marginal damages}\\ \text{to output and utility}}} + \underbrace{\sum_{j=0}^{\infty} \beta_G^j \frac{\partial S_{t+j}}{\partial E_t}\left[\frac{-\partial F_{Kt+j}}{\partial S_{t+j}}\frac{\Phi_{t-1+j}}{\lambda_t \beta_G}\right]}_{\substack{\text{Marginal damages to}\\ \text{to marginal product of capital}}} + \underbrace{\left(\frac{-\partial F_{Kt}}{\partial E_t}\right)\frac{\Phi_{t-1}}{\lambda_t \beta_G}}_{\substack{\text{Marginal energy use impact}\\ \text{on marginal product of capital}}} \tag{22}$$

[5] Non-separability specifically implies that the planner must account for the effects of climate change on the household's intertemporal marginal rate of substitution if the shadow value of capital subsidies is changing over time. Both for parsimony and given the limited empirical evidence on how direct utility impacts of climate change (e.g., damages to cultural heritage sites) enter households' preferences, Proposition 1 thus focuses on the benchmark case with separability.

where λ_t and Φ_t denote the planner's Lagrange multipliers on the aggregate resource constraint Eq. (11) and the no capital income subsidy constraint Eq. (21), respectively.

Proof. See Appendix A. The central implication of Proposition 1 is that a planner endorsing social discounting may not actually want to impose a Pigouvian carbon tax equal to the socially discounted marginal damages of emissions if he cannot subsidize capital income at the same time. Intuitively, this is because the planner cares about the overall level of assets given to future generations – both natural and man-made – and consequently takes into account the general equilibrium effects of carbon taxes on households' savings behavior. Without a capital income subsidy, private savings are too low. Consequently, climate policy's impacts on private investment can now have first-order welfare effects, analogous to settings with pre-existing distortions from income taxes (e.g., Bovenberg and Goulder, 1996) or imperfect competition (e.g., Ryan, 2012). On the one hand, carbon taxes should be increased (all else equal) to the extent that climate change will decrease the marginal product of capital. Mathematically, this can be seen from Eq. (22) by noting that, if $\beta_G > \beta_H$, the no-subsidy constraint will be binding ($\Phi_t > 0$), so that a negative effect of carbon stocks on the marginal product of capital $\left(\frac{\partial F_{Kt+j}}{\partial S_{t+j}} < 0\right)$ yields a positive overall term for the social value of marginal damages to the marginal product of capital in Eq. (22). Intuitively, this is because these damages reduce private returns to savings even further away from the social optimum. On the other hand, carbon taxes should be reduced (all else equal) to the extent that reductions in energy usage lower the marginal product of capital. Mathematically, this follows readily from Eq. (22) by analogous logic when $\frac{\partial F_{Kt}}{\partial E_t} > 0$. Intuitively, this is again because households are already insufficiently incentivized to save, implying that any further decrease in the marginal product of capital due to higher energy costs will exacerbate this social inefficiency. If energy and capital are sufficiently complementary in final goods production, the constrained-optimal carbon tax may thus be lower than the social cost of carbon (i.e., the SDV of marginal damages to output and utility).

4. Distortionary fiscal setting

The analysis presented so far has assumed that governments can implement lump-sum transfers. In reality, however, public funds must be raised through distortionary instruments, such as labor income taxes. This section thus adds government revenue requirements and linear tax instruments in the Ramsey tradition (see, e.g., Chari and Kehoe, 1999) to the model. The objectives of this extension are to study both (i) the robustness of the desirability of capital income subsidies to a setting where they must be financed with distortionary taxes, and (ii) the broader fiscal policy implications of differential social discounting. First, I introduce preferences for leisure, changing household utility to:

$$U_0 \equiv \sum_{t=0}^{\infty} \beta_H^t U(C_t, L_t, S_t). \tag{23}$$

Second, the household may now be subject to labor and consumption taxes in addition to capital income taxes:

$$(1+\tau_{ct})C_t + \rho_t B_{t+1} + K_{t+1} \leq w_t(1-\tau_{lt})L_t + \{1 + (r_t - \delta)(1-\tau_{kt})\}K_t + B_t + \Pi_t. \tag{24}$$

Here, τ_{ct} denotes the consumption tax in period t, B_{t+1} is purchases of one-period government bonds at price ρ_t, K_{t+1} denotes the household's capital holdings in period $t+1$, τ_{lt} the linear labor income tax, τ_{kt} the linear net-of-depreciation capital income tax, B_t repayments of government bond holdings, and the remaining variables are as defined above. The household's optimality conditions for its labor-consumption and savings decisions in this setup become:

$$\frac{-U_{lt}}{U_{ct}} = w_t \frac{(1-\tau_{lt})}{(1+\tau_{ct})} \tag{25}$$

$$\frac{U_{ct}}{U_{ct+1}\beta_H} = \{1 + (r_{t+1} - \delta)(1-\tau_{kt+1})\}\frac{(1+\tau_{ct})}{(1+\tau_{ct+1})}. \tag{26}$$

The government must finance an exogenously given sequence of revenue requirements $\{G_t\}_{t=0}^{\infty}$ assumed to be strictly positive for at least one t and sufficiently large so as to necessitate distortionary taxation. That is, the revenues from the exogenously given initial period tax rates ($\overline{\tau_{k0}}$, $\overline{\tau_{c0}}$) and optimized emissions taxes are assumed to be insufficient to cover the present value of all government spending requirements. As is common, G_t is moreover modeled as wasteful government consumption. The quantitative version of the model further adds social transfers to households. The government can raise revenues by issuing bonds and levying taxes, with corresponding flow budget constraint:

$$G_t + B_t = \tau_{lt} w_t L_t + \tau_{ct} C_t + \tau_{kt}(r_t - \delta)K_t + \tau_{Et} E_t + \rho_t B_{t+1}. \tag{27}$$

It should be noted that not all of these tax instruments are needed to form a "complete" tax system (Chari and Kehoe, 1999). That is, the same allocation can be decentralized by many different tax systems that correspond to the same overall *wedges* between the relevant marginal rates of substitution and transformation. For example, define the following term:

$$\text{Effective labor-consumption wedge} \quad \tau_{lct} \equiv 1 - \frac{(1-\tau_{lt})}{(1+\tau_{ct})}.$$

It is clear from Eq. (25) that many different labor-consumption tax combinations map into the same effective wedge. Below, the results are thus presented first in terms of optimal wedges, and then highlighting two examples of decentralizations through specific tax instruments.

Competitive equilibrium in this economy is defined in the standard way, with the addition of the carbon cycle constraint and pre-industrial concentrations S_0 as initial condition. The government's objective is to implement the competitive equilibrium that yields the highest household lifetime utility – discounted at the social rate β_G – for a given set of initial conditions (K_0, B_0, S_0, $\overline{\tau_{k0}}$, $\overline{\tau_{c0}}$). As is standard, the initial capital income tax $\overline{\tau_{k0}}$ and consumption tax $\overline{\tau_{c0}}$ are assumed to be exogenously given as they can otherwise be used as effective lump-sum taxes. I also assume that the government can commit to a sequence of taxes.[6]

The set of allocations that can be decentralized as a competitive equilibrium can be characterized by two sets of constraints: feasibility and an "implementability" constraint (Eq. (IMP)) that captures the optimizing behavior of households and firms. In the present setting, it is straightforward to show (see Chari and Kehoe, 1999, for a general discussion, and Barrage, 2016, for a derivation in a climate-economy model similar to the present framework) that the IMP is:

$$\sum_{t=0}^{\infty} \beta_H^t \underbrace{(U_{ct}C_t + U_{lt}L_t)}_{\equiv W_t} = \left\{\frac{U_{c0}}{1+\overline{\tau_{c0}}}[K_0\{1 + (F_{K0} - \delta)(1-\overline{\tau_{k0}})\}]\right\} \tag{IMP}$$

The main difference between Eq. (IMP) and the standard setting is that one must be careful to employ the household's discount factor β_H rather than the planner's β_G in the derivation of Eq. (IMP) as it captures the optimizing behavior of households.

Let ϕ denote the Lagrange multiplier on Eq. (IMP) in the planner's problem, and let $W_t \equiv (U_{ct}C_t + U_{lt}L_t)$ denote the bracketed term on the left-hand side of Eq. (IMP). It is straightforward to show that the planner's first-order conditions (for $t > 0$)[7] imply the following intra-

[6] This assumption is common in the Ramsey taxation literature, but is known not to be without loss of generality (e.g., Klein and Ríos-Rull, 2003; Benhabib and Rustichini, 1997). However, focusing on carbon taxes alongside other distortionary taxes, Schmitt (2014) finds that relaxing the assumption of commitment has only a minor quantitative effect on optimal policy.

[7] The restriction to $t > 0$ stems from the fact that the planner's first-order conditions differ at $t = 0$ versus all subsequent periods. This is because the $t = 0$ conditions take into

and intertemporal optimality conditions, respectively:

$$\frac{-U_{lt}}{U_{ct}} = F_{Lt}\left(1 + \phi\left[\frac{W_{ct}}{U_{ct}} + \frac{W_{lt}}{U_{ct}F_{Lt}}\right]\left(\frac{\beta_H}{\beta_G}\right)^t\right) \tag{28}$$

$$\frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^t U_{ct} + \phi W_{ct}\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} U_{ct+1} + \phi W_{ct+1}\right]}\frac{1}{\beta_H} = [(1-\delta) + F_{Kt+1}] \tag{29}$$

where W_{ct} (W_{lt}) denotes the partial derivative of expression W_t in Eq. (IMP) with respect to consumption (labor) at time t. As is well-known, the precise nature of optimal policy thus depends on the structure of preferences. I proceed focusing on a commonly used constant elasticity of substitution specification that satisfies consistency with balanced growth (see King et al., 2001), and present results for additional utility functions in the numerical analysis below. Assume that preferences are separable and of the form:

$$U(C_t, L_t, S_t) = \log C_t + v(L_t) + \vartheta(S_t) \tag{A}$$

where $v(L_t)$ is an increasing and concave function of leisure. With preferences of the form Eq. (A), it is straightforward to show that $W_{ct} = 0$ for $t > 0$. Consequently, the planner's optimality condition for investment (Eq. (29)) becomes:

$$\frac{U_{ct}}{U_{ct+1}\beta_H} = [(1-\delta) + F_{Kt+1}]\left(\frac{\beta_G}{\beta_H}\right). \tag{30}$$

Condition Eq. (30) immediately reveals that the planner wants the household's intertemporal marginal rate of substitution (MRS) to exceed the intertemporal marginal rate of transformation (MRT) when $\beta_G > \beta_H$. In order to decentralize this allocation, on net the optimal tax policy thus needs to create an effective subsidy on the return to capital.

Next, in order to characterize intratemporal wedges, assume the following form for $v(L_t)$:

$$v(L_t) = -v_1 L_t^{v_2}. \tag{A'}$$

Specification Eq. (A') is increasing and concave in leisure as long as $v_1 > 0$ and $v_2 > 1$. Given Eq. (A'), it is straightforward to show that $W_{lt} = U_{lt}v_2$. Substituting into the planner's intratemporal optimality condition Eq. (28) and rearranging then yields (for $t > 0$):

$$\frac{-U_{lt}}{U_{ct}} = F_{Lt}\left[1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t\right]^{-1}. \tag{31}$$

Comparing Eq. (31) to the household's intratemporal optimality condition Eq. (25) reveals that the optimal intratemporal distortion (i.e., wedge distorting the labor-consumption tradeoff) is *decreasing over time* when $\beta_G > \beta_H$. Proposition 2 summarizes these results formally, and Corollary 1 characterizes specific combinations of tax instruments that can decentralize the optimal allocation.

Proposition 2. *Assume that the planner values the future more than households ($\beta_G > \beta_H$) and that emissions taxes are set optimally.*

- *If preferences satisfy* Eq.(A)*, the optimal fiscal policy also creates an effective capital income subsidy for all $t > 1$ (i.e., a negative wedge between the intertemporal marginal rates of transformation and substitution).*
- *If preferences satisfy* Eqs.(A) *and*(A')*, the optimal fiscal policy creates an effective labor-consumption tax wedge that is decreasing over time for all $t > 1$.*

Proof. See Appendix B.

(*footnote continued*)
account the effect of initial allocations on the initial value of the household's assets in Eq. (IMP).

Corollary 1. *Assume that the planner values the future more than households ($\beta_G > \beta_H$) and that preferences satisfy Eqs.*(A) *and*(A')*.* The optimal allocation can then be decentralized by a tax system consisting of the following (for $t > 1$):

- *An emissions tax equal to the socially discounted marginal damages of carbon emissions,*
- *A capital income subsidy,*
- *Plus either one of the following:*
  - *A labor income tax that is weakly positive but decreasing over time, or:*
  - *A consumption tax that is weakly positive but decreasing over time.*

  In contrast, if the planner adopts the household's pure rate of social time preference ($\beta_G = \beta_H$) and preferences satisfy (A) and (A'), then the optimal tax system consists of the following (for $t > 1$):
- *An emissions tax equal to the socially discounted marginal damages of carbon emissions,*
- *Zero capital income taxes or subsidies,*
- *Plus either one of the following:*
  - *A positive and constant labor income tax, or:*
  - *A positive and constant consumption tax*

Proof. See Appendix B.

It is useful to note that the tax policies described in Corollary 1 are only examples, and that the properties of individual taxes may differ in alternative decentralizations. For example, it is straightforward to show that the optimal allocation can also be decentralized by emissions taxes coupled with decreasing consumption taxes and increasing labor income taxes. Intuitively, the decline in consumption taxes serves to increase the returns to savings and implements the effective intertemporal subsidy as per Eq. (26). If this rate of decline in consumption taxes is faster than the optimal decline in the effective labor-consumption wedge, however, labor taxes must increase over time in order to compensate. On net, however, the optimal effective labor-consumption tax wedge is still decreasing over time as per Proposition 2.

The central implication of these theoretical results is thus that differential discounting fundamentally alters optimal fiscal policy prescriptions in the Ramsey framework. With standard discounting ($\beta_G = \beta_H$), optimal policy features zero capital income taxes, in line with the classic result that intertemporal distortions are undesirable in a wide range of settings (see, e.g., Judd, 1985, 1999; Atkeson et al., 1999; Acemoglu et al., 2011). Standard discounting also implies that intratemporal distortions be smoothed over time, again in line with certain classic results from the literature (e.g., the uniform commodity taxation result applied to consumption over time as discussed in Chari and Kehoe, 1999). In contrast, with differential social discounting, the optimal tax code induces an intertemporal wedge to effectively subsidize capital income, and features intratemporal distortions that are decreasing over time. To the best of my knowledge, these implications of differential discounting have not been previously formalized in this setting.

5. Quantitative analysis

5.1. Calibration

This section assesses the quantitative importance of differential social discounting by extending two dynamic general equilibrium climate-economy model calibrations of the presented framework:

5.1.1. GHKT-B

First, in order to match the benchmark theoretical model, I build on the generalized numerical implementation of Golosov, Hassler, Krusell, and Tsyvinski's model (2014) (GHKT) by Barrage (2014), which allows for general CRRA preferences $\left(U(C_t) = \frac{C_t^{1-\sigma}-1}{1-\sigma}\right)$. The application adds three new elements:(i) differential discounting, (ii) capital income tax

constraints, and (iii) a nested CES production structure to consider capital-energy elasticities of substitution below unity. In contrast, the GHKT framework focuses on Cobb-Douglas technology. Based on a review of the literature by Van der Werf (2008), I consider the following structure:

$$Y_t = (1 - D(S_t))\cdot A_t \left[\varkappa [K_t^{\alpha} L_t^{1-\alpha}]^{\frac{\varepsilon-1}{\varepsilon}} + (1-\varkappa)[E_t]^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε denotes the elasticity of substitution between energy inputs E_t and the capital-labor aggregate. In addition to these changes, I also simplify the energy sector representation compared to GHKT. Appendix C describes these changes and calibration in detail.

With regard to the capital income tax constraints, the quantitative analysis considers two cases. One incorporates the no-subsidy constraint Eq. (21) analyzed in Section 3. The other considers the possibility that the planner cannot reduce capital income taxes below some positive level $\tau_{kt+1} \geq \overline{\tau_k} > 0$. The assumption of a minimum politically acceptable capital income tax $\tau_{kt+1} \geq \overline{\tau_k}$ corresponds to the primal constraint:

$$\frac{U_{ct}}{\beta_H U_{ct+1}} \leq \{1 + (F_{Kt+1} - \delta)(1 - \overline{\tau_k})\}. \tag{32}$$

As shown in the Online Appendix, the constrained-optimal carbon tax subject to Eq. (32) is structurally the same as the benchmark expression in Proposition 1 except that both terms relating to the capital income tax constraint are now additionally multiplied by $(1 - \overline{\tau_k})$. Of course, the optimal allocation at which the expression will be evaluated, and consequently the optimal carbon tax level, may be substantially different, as the numerical results will show.

5.1.2. COMET

Second, in order to quantify optimal policies in the extended theoretical framework with endogenous labor supply and distortionary taxes, I integrate differential discounting into the COMET model (Barrage, 2016). The COMET builds upon the model structure presented above and on the seminal DICE framework of Nordhaus (see, e.g., Nordhaus, 2008, 2010a,b), which is one of three models currently used by the U.S. Government to value the social cost of carbon (Interagency Working Group, 2010). The COMET expands upon DICE in several ways, including preferences for leisure and the climate, energy production, tax policy choice, and government spending requirements (see Barrage, 2016). Household utility in the COMET is given by:

$$U(C_t, L_t, T_t) = \left\{ \frac{[C_t \cdot (1 - \varsigma L_t)^{\gamma}]^{1-\sigma}}{1-\sigma} \right\} + \frac{(1 + \alpha_0 T_t^2)^{-(1-\sigma)}}{1-\sigma}. \tag{33}$$

The parameter ς is introduced in order to ensure that the calibration can simultaneously match (i) a desired intertemporal elasticity of substitution, (ii) a Frisch elasticity of labor supply of 0.78, and (iii) and to rationalize base year (2005) labor supply as estimated from OECD data (see Barrage, 2016 for details). The variable T_t denotes mean atmospheric surface temperature change over pre-industrial levels. Government revenue requirements and expenditure patterns are calibrated based on IMF Government Finance Statistics.[8] On the macroeconomic side, the COMET further adopts the productivity, clean energy technology, and population growth rate projections of the DICE/RICE model family (Nordhaus, 2008, 2010a,b).

5.1.3. Benchmark

Both models assume a common set of key parameter values for the benchmark calibration. These are summarized in Table 1.

Table 1
Benchmark case — key parameters.

Parameter	Value	Source
σ	1.5	Match Nordhaus (2008)
ε	1	GHKT, DICE
β_H	0.985	Nordhaus (2008)
β_G	0.999	Stern (2006)

5.2. Optimal policy results

5.2.1. First-best benchmark(GHKT-B)

Fig. 1 displays optimal carbon taxes in the benchmark first-best setting. As expected, climate policy is significantly more stringent with social discounting (blue line with hexagons) than when the planner adopts households' pure rate of time preference (black line with asterisks). The estimated optimal tax for 2020 more than doubles from $45 ($2015/mtC) to $94. Fig. 2 showcases the accompanying optimal capital income tax rates in each scenario. Implementing the optimal allocation with social discounting requires large capital income subsidies of around 50%. Table 2 summarizes these results and shows that capital income subsidies remain high across alternative values of the consumption elasticity.

5.2.2. Constrained-optimal climate policy(GHKT-B)

Table 3 compares optimal carbon prices in the first-best setting with optimal levies when the planner is constrained either from subsidizing capital income Eq. (21) or from lowering capital income taxes below $\overline{\tau_k} = 30\%$ (Eq. (32)). While the effect of the no-subsidy constraint on carbon taxes is theoretically ambiguous, the quantitative results suggest a negative effect for all parameters considered. If the planner cannot subsidize capital income, he should adjust carbon taxes downward by 5–30 %. With a minimum capital income tax of $\overline{\tau_k} = 30\%$, optimal carbon tax adjustments are larger, ranging from −10% to −45%. As predicted by the theoretical results, climate policy is adjusted downward more strongly when capital and energy are more complementary in production, as this implies a larger negative effect of carbon taxes on the returns to saving (by making energy inputs more expensive). The constrained-optimal policy thus strikes a balance between incentivizing investments in environmental and man-made assets for the overall benefit of future generations. Table 4 shows that these results are similar for alternative values of the consumption elasticity and capital-energy elasticity of substitution.

5.2.3. Distortionary fiscal setting(COMET)

This section evaluates the quantitative implications of social discounting in a Ramsey second-best fiscal setting. While the theoretical results reveal that capital income subsidies remain desirable even when they must be financed with distortionary taxes, the empirical magnitude of this and other tax policy deviations from standard recommendations remains an open question. Figs. 3 and 4 display optimal carbon taxes and temperature change. As expected, social discounting leads to significantly larger carbon taxes and reduces optimal peak global temperature change to below 2 ° C. In contrast, with standard discounting, the optimal policy limits climate change to around 3 ° C, in line with the DICE/RICE model family (Nordhaus, 2008, 2010a,b, etc.) (Table 5).

While the climate policy implications of social discounting are well-known, Figs. 5 and 6 and Table 5 showcase the broader fiscal policy adjustments that this approach also requires, focusing on a decentralization with capital income and labor taxes. These adjustments are large for the discount factors advocated in the literature: For labor taxes, social discounting implies that labor should first be taxed at a high rate of 53%, followed by a continual decline to 33% by the end of

[8] In the model base year 2005, the PPP-adjusted GDP-weighted average share of government expenditures is ~ 33% of GDP. Expenditures are further broken down into government consumption ($G_t^C \sim 57\%$) and social transfers ($G_t^T \sim 43\%$). In line with other fiscal computable general equilibrium and optimization models (e.g., Jones et al., 1993; Goulder, 1995), government expenditure G_t grows at the rates of labor productivity and population growth, with G_t^C and G_t^T evolving at constant shares proportional to G_t.

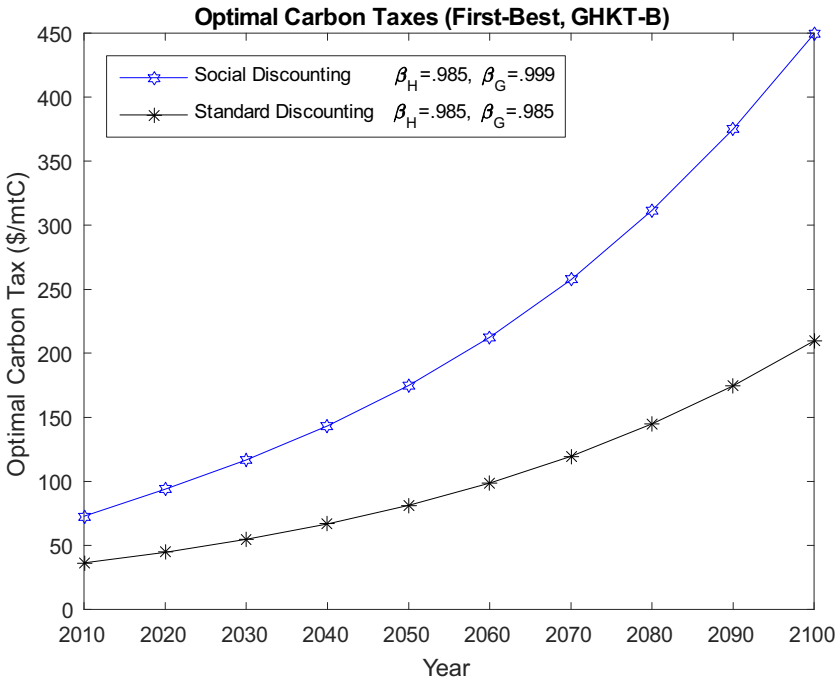

Fig. 1. First-best carbon taxes (GHKT-B). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

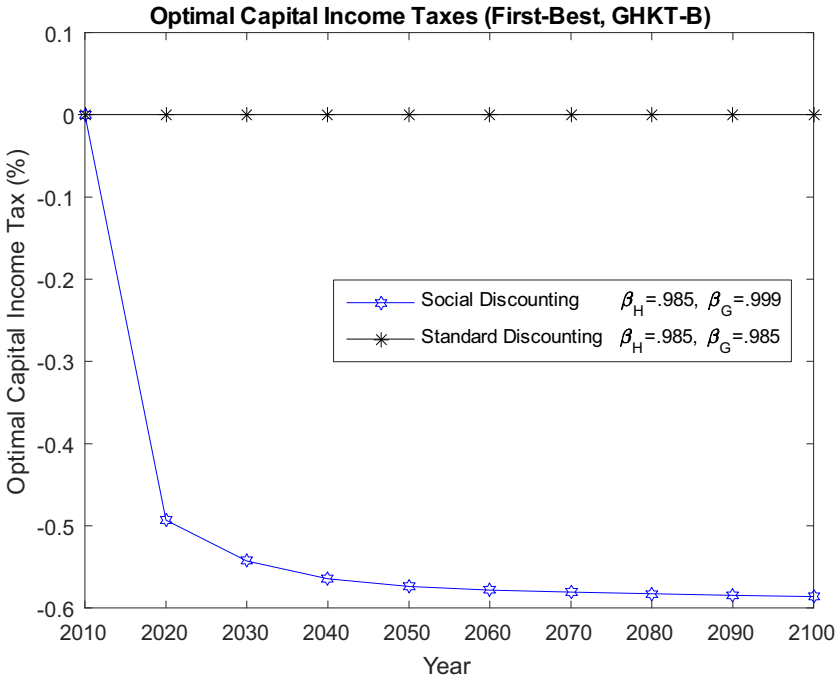

Fig. 2. First-best capital income taxes (GHKT-B).

the century. Decentralizing the optimal allocation further requires capital income subsidies ranging from 29% to 74% by the end of the century. In contrast, with standard discounting, the model yields the standard results that labor taxes should be constant at 41%, and that capital income taxes converge to zero after the first period.[9]

Table 6 showcases the sensitivity of optimal fiscal policies for

Table 2
Sensitivity of first-best policies to consumption elasticity (2020).

	Discounting			
Consumption elasticity	Standard ($\beta_G = \beta_H = 0.985$)		($\beta_G = .999$; $\beta_H = 0.985$)	
	Carbon (τ^*_{Et})	Capital Inc. (τ^*_{kt+1})	Carbon (τ^*_{Et})	Capital Inc. (τ^*_{kt+1})
$\sigma = 1$	\$70	0%	\$639	−62%
$\sigma = 1.5$	\$45	0%	\$94	−49%
$\sigma = 2$	\$34	0%	\$53	−43%

Optimal carbon taxes in \$2015 per metric ton of carbon (mtC).

[9] Capital income taxes do not immediately jump to zero here after $t = 1$ because of the nature of the non-separable utility function (33) in the COMET, for which capital income taxes go to zero only as labor supply converges to a constant level. In contrast, with utility function (A'), the optimal capital income tax as of $t = 1$ would be exactly equal to zero.

Table 3
Constrained-optimal carbon taxes.

	Capital income tax scenario				
	Optimized	Constrained		Constrained	
	$\tau_{kt+1} \approx -80\%$	$\tau_{kt+1} = 0$		$\tau_{kt+1} = 30\%$	
Energy-KL elasticity	Level	Level	Avg. adjustment	Level	Avg. adjustment
	($/mtC, 2020)	($/mtC, 2020)	(%, 2020–2100)	($/mtC, 2020)	(%, 2020–2100)
$\varepsilon = 1$	$94	$90	−5%	$87	−11%
$\varepsilon = 0.4$	$95	$65	−27%	$46	−44%

Table 4
Sensitivity of constrained-optimal carbon taxes.

	Capital income tax scenario				
	Optimized	Constrained		Constrained	
	$\tau_{kt+1} \approx -80\%$	$\tau_{kt+1} = 0$		$\tau_{kt+1} = 30\%$	
Energy-KL elasticity	Level	Level	Avg. adjustment	Level	Avg. adjustment
	($/mtC, 2020)	($/mtC, 2020)	(%, 2020–2100)	($/mtC, 2020)	(%, 2020–2100)
Consumption elasticity $\sigma = 1$					
$\varepsilon = 1$	$639	$612	−6%	$588	−10%
$\varepsilon = 0.4$	$609	$556	−11%	$504	−20%
Consumption elasticity $\sigma = 2$					
$\varepsilon = 1$	$53	$51	−5%	$49	−11%
$\varepsilon = 0.4$	$53	$30	−37%	$14	−63%

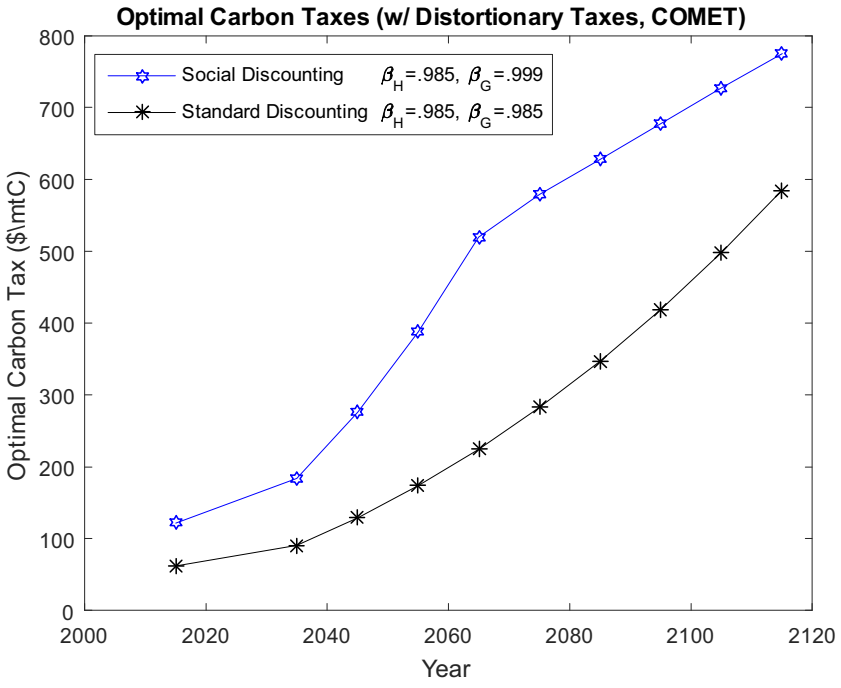

Fig. 3. Optimal carbon taxes in second-best fiscal setting.

alternative values of the consumption elasticity and an alternative Frisch elasticity of labor supply of 2.84 (Chetty et al., 2011). The results continue to show that social discounting implies a quantitatively significant departure from both current fiscal policy practice and standard recommendations in the Ramsey framework.

5.2.4. Welfare

One of the core results of this paper is that, in general equilibrium, a planner's decision to weigh future generations' welfare at a higher rate than households has fundamental implications not only for carbon pricing but for fiscal policy more broadly. The core intuition is that climate protection is only one of several channels through which the

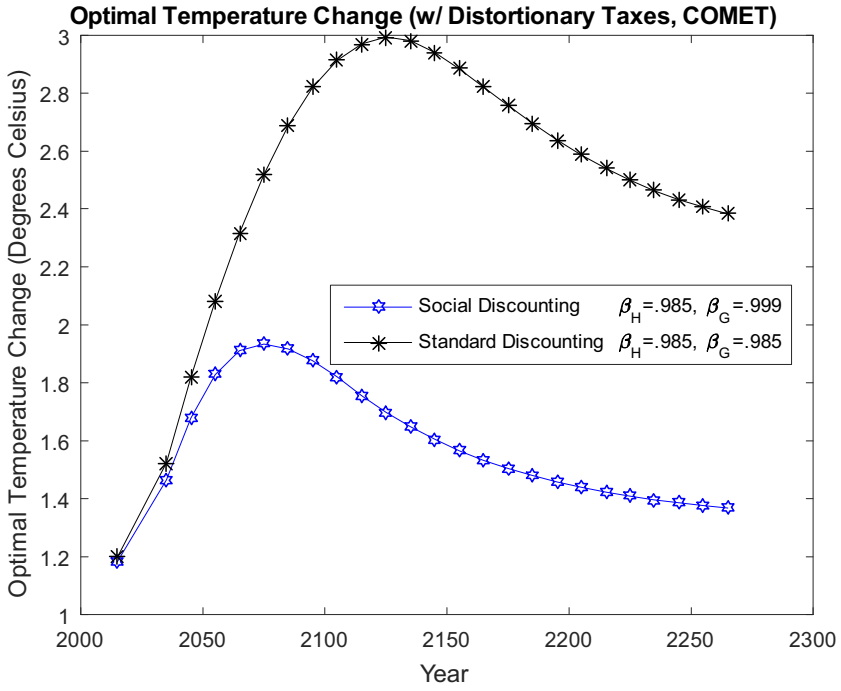
Optimal Temperature Change (w/ Distortionary Taxes, COMET)
Optimal Temperature Change (Degrees Celsius)
Year
Social Discounting β_H=.985, β_G=.999
Standard Discounting β_H=.985, β_G=.985

Fig. 4. Temperature change in second-best fiscal setting.

Table 5
Optimal policy in second-best fiscal setting.

	Discounting			
	Standard ($\beta_G = \beta_H = 0.985$)		Social ($\beta_G = 0.999; \beta_H = 0.985$)	
Tax	2025	2100	2025	2100
Carbon (τ^*_{Et})	$90	$585	$183	$775
Capital Inc. (τ^*_{kt+1})	2%	0%	− 29%	− 74%
Labor Inc. (τ^*_{lt})	41%	41%	53%	33%

planner can make future generations better off. In order to formalize this point further, this section compares the welfare benefits of competing policy adjustments that a socially discounting planner would like to make.

The baseline scenario against which policy changes are compared assumes that all taxes are set at levels that would be optimal with standard discounting ($\tau_{Et}^{\beta_G=.985}$, $\tau_{kt+1}^{\beta_G=.985}$, $\tau_{lt}^{\beta_G=.985}$). The simulations then allow the socially discounting planner to optimize one policy lever at a time. For example, in the GHKT-B model, the planner gets to impose constrained-optimal socially discounted carbon taxes $\{\tau^*_{Et}\}$ while capital income taxes are restricted to be zero (i.e., at their first-best level with

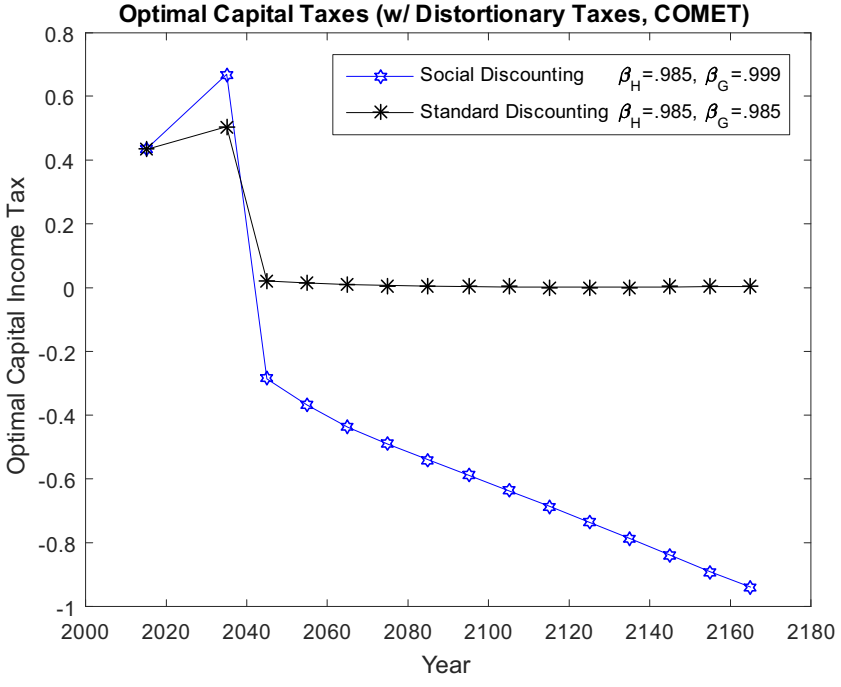
Optimal Capital Taxes (w/ Distortionary Taxes, COMET)
Optimal Capital Income Tax
Year
Social Discounting β_H=.985, β_G=.999
Standard Discounting β_H=.985, β_G=.985

Fig. 5. Optimal capital taxes in second-best fiscal setting.

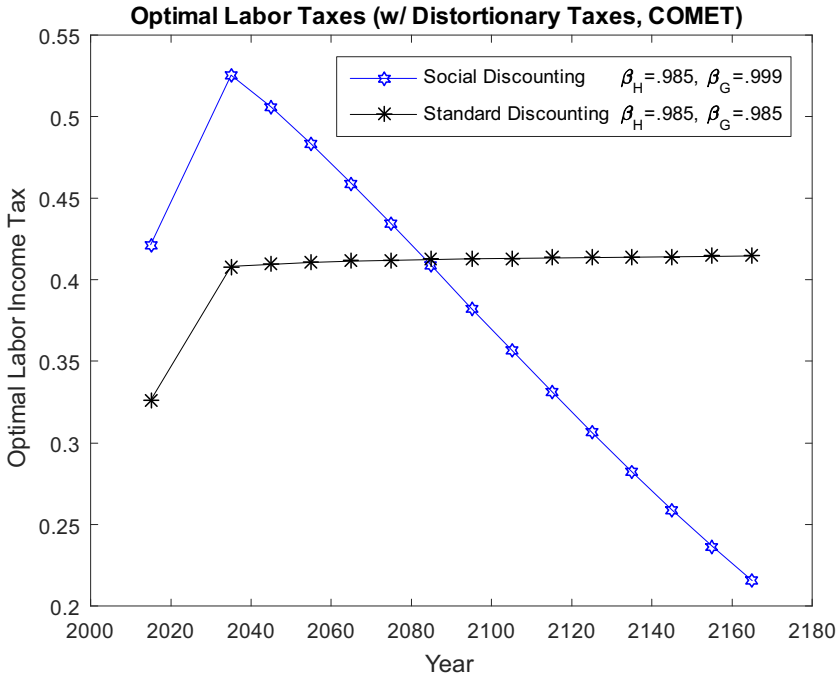

Fig. 6. Optimal labor taxes in second-best fiscal setting.

Table 6
Sensitivity of optimal policy in second-best fiscal setting.

	Discounting			
Tax	Standard ($\beta_G = \beta_H = 0.985$)		Social ($\beta_G = 0.999$; $\beta_H = 0.985$)	
	2025	2105	2025	2105
Frisch labor supply elasticity = 2.84				
Carbon (τ^*_{Et})	$82	$519	$133	$742
Capital Inc. (τ^*_{kt+1})	5%	0%	−12%	−59%
Labor Inc. (τ^*_{lt})	42%	44%	53%	38%
Consumption elasticity $\sigma = 1.1$				
Carbon (τ^*_{Et})	$136	$663	$310	$795
Capital Inc. (τ^*_{kt+1})	2%	0%	−32%	−86%
Labor Inc. (τ^*_{lt})	39%	39%	54%	35%
Consumption elasticity $\sigma = 2$				
Carbon (τ^*_{Et})	$58	$528	$110	$749
Capital Inc. (τ^*_{kt+1})	2%	0%	−28%	−59%
Labor Inc. (τ^*_{lt})	43%	44%	52%	32%

standard discounting $\tau^{\beta_G=.985}_{kt+1}$). Importantly, in all of these simulations, it is the relevant intertemporal or intratemporal *wedges* that are restricted to match a given tax level, so that the planner can only adjust the policy margin in question (e.g., he cannot introduce a consumption tax that mimics a capital income tax in addition to setting τ^*_{Et} in the aforementioned example). Conceptually, the simulations are meant to assess which policy adjustments should be a priority for a planner who weighs the future at a higher rate than households. The literature has focused heavily on climate policy adjustments, particularly whether the social cost of carbon should be calculated based on standard (Nordhaus, 2010a,b) or social (Stern, 2006) discounting. However, the results suggest that other policy adjustments may, in some cases, yield higher welfare benefits, indicating that a broader policy and modeling focus may be of value to the literature concerned about intergenerational equity.

Before discussing the results in detail, two notes are in order. First, in the model with distortionary taxes (COMET), labor taxes are generally restricted to be *constant* after the first period, rather than equal to their (constant) benchmark level ($\tau^{\beta_G=.985}_{lt}$). Intuitively, this is to provide an additional degree of freedom to the planner in meeting his intertemporal government budget constraint (which includes a revenue requirement). That is, without this adjustment, the planner would have a very limited ability to change any one policy lever in the desirable direction for social discounting (e.g., to introduce capital income subsidies) while meeting his revenue requirement in competitive equilibrium. Consequently, in order to improve comparability across the two models, I introduce this additional degree of freedom for all but the benchmark scenarios.

Second, for completeness, I also run each model with a 'no carbon tax' scenario as an alternative baseline. However, it is important to note the limitations of these scenarios. First, the GHKT-B model assumes that the carbon emissions intensity of fossil fuels exogenously declines over time, leveling off to zero within two centuries. Consequently, even the 'no carbon tax' scenario does not reflect a case with unlimited climate change. Second, the smooth damage function is also arguably not ideally suited to study the welfare consequences of climate change above certain levels where, e.g., tipping points begin to play an important role (see, e.g., Lemoine and Traeger, 2014). The same concern applies to the COMET, which builds on the seminal DICE model's (Nordhaus, 2008) convex damage function formulation. While the welfare consequences of 'no carbon tax' may thus be underestimated, importantly, this is not the focus on the present analysis, which centers on the welfare effects of policy adjustments at stake in the debate and range between standard and social discounting.

Table 7 displays the results from the GHKT-B model. In the benchmark calibration, increasing carbon taxes from standard to socially discounted levels yields a welfare gain equivalent to a permanent 0.17% consumption increase. Introducing capital income subsidies yields a significantly higher welfare gain of +0.43%. Both policies together yield a combined benefit of +0.63%. The biggest welfare gain, however, stems from having a carbon tax at all: even a standard-discounting (e.g., Nordhaus, 2008) carbon tax provides a welfare gain equivalent to a permanent 1.06% consumption increase. This pattern

Table 7
Welfare effects for socially discounting planner (GHKT-B).

Tax policy scenario:		Δ Welfare[a]			
Carbon	Capital	Benchmark	$\varepsilon = 0.4$	$\sigma = 1$	$\sigma = 2$
1. $\tau_{Et}^{\beta_G=.985}$	$\tau_{kt+1}^{\beta_G=.985}$ (None)	–	–	–	–
2. τ_{Et}^{*}	$\tau_{kt+1}^{\beta_G=.985}$ (None)	+0.17%	+0.013%	+2.25%	+0.022%
3. $\tau_{Et}^{\beta_G=.985}$	τ_{kt+1}^{*}	+0.43%	+0.42%	+0.48%	+0.37%
4. τ_{Et}^{*}	τ_{kt+1}^{*}	+0.62%	+0.46%	+2.82%	+0.40%
5. None	None	−1.06%	−0.11%	−4.01%	−0.28%

[a] Equivalent variation permanent percentage change in consumption.

changes, however, for alternative parameter values. With both CES production and a consumption elasticity of $\sigma = 2$, capital income subsidies provide the highest welfare improvement. In contrast, with logarithmic preferences, the welfare gains from adjusting carbon taxes to socially discounted levels exceed those of capital income subsidies by more than a factor of four.

Table 8 displays the results from the COMET model. Once again, the welfare benefits from increasing carbon taxes (+0.96%) are lower than those from adjusting either capital income taxes (+1.36%) or labor income taxes (+2.30%) at the benchmark. At the same time, all are eclipsed by the benefits of even a standard-discounted carbon tax (+4.61%). In alternative calibrations, however, the pattern again switches: labor income tax adjustments yield the highest welfare benefits among the policies with both a higher consumption elasticity ($\sigma = 2$) or a higher Frisch elasticity of labor supply (2.84, Chetty et al., 2011). In contrast, with a lower consumption elasticity, the relative importance of carbon taxes and adjustments again increases.

Overall, the results from both models thus indicate that the relative welfare ranking of adjusting different policies for the benefit of future generations can vary depending on the parameter values. While the calculations are subject to several important limitations (e.g., smooth climate damage functions), they do again illustrate the core point that a singular focus on stricter climate policy may not be the optimal response to increased concern for future generations.

6. Conclusion

The rate at which policy-makers should discount the future has long been one of the most intensely debated questions in public economics. This issue is moreover central to current debates on climate policy, where a fundamental disagreement has emerged between those calling for an ethics-based calibration of the pure rate of social time preference (e.g., Stern, 2006), and those advocating that discounting parameters should be calibrated to match households' revealed time preferences and market interest rates (e.g., Nordhaus, 2007). In order to reconcile these perspectives, a number of authors have proposed allowing the planner to select a different utility discount rate than exhibited by households (e.g., Goulder and Williams, 2012). This approach can moreover be explicitly microfounded by a planner who places higher welfare weights on future generations than currently alive households in an overlapping generations framework (e.g., Farhi and Werning, 2005), showcasing its potential to address the intergenerational equity concerns at the heart of the debate.

This paper formalizes the broader policy implications of differential discounting in a dynamic general equilibrium climate-economy model with an extension to fiscal policy in the Ramsey tradition. I quantify the magnitude of the required policy changes and their welfare effects by building on the integrated assessment models of both GHKT (2014), and an extended version of the seminal DICE model (Nordhaus, 2008) that incorporates fiscal policy and endogenizes key variables such as labor supply (COMET, Barrage, 2016).

The main result is that a policy-maker's choice to adopt differential social discounting fundamentally alters prescriptions for both environmental and fiscal policy. On the one hand, decentralizing the optimal allocation would require radical changes to tax policy, including a transition to substantial capital income subsidies (~ 30–75 %), and, depending on the fiscal setting, a change to intratemporal distortions that decrease over time, such as labor income taxes that decline from 53% to 33% by the end of the century. On the other hand, if governments cannot subsidize capital income, the constrained-optimal carbon tax may no longer equal the social cost of carbon (the Pigouvian tax), and must be adjusted for the general equilibrium effects of climate policy on households' incentives to save. These effects depend on the precise nature of climate damages and on the complementarity between energy and capital in production, highlighting the value of future research on both of these parameters. Quantitatively, I find that the constrained-optimal carbon tax is 5–60 % below the first-best levy, indicating that even a planner with 'Sternian' social discounting preferences may not want to impose a 'Sternian' carbon tax due to its effects on the overall wealth that will be provided to future generations. The welfare calculations similarly suggest that, for certain areas of the parameter space, fiscal policy reform may provide a better way of redistributing to the future than tighter climate policy.

One important objection to these results is that they fail to capture the potentially limited substitutability between the climate and man-made assets and goods, which can increase the optimal carbon price as much as social discounting (Sterner and Persson, 2008; Traeger, 2011). Formally accounting for this limited substitutability would moreover not lead to the surprising broader policy implications of social discounting outlined in this paper.

In general, it is essential to note that the results of this study pertain to a disconnect between the planner's and the household's utility discount factors based on an ethical disagreement about the value of future generations' welfare. Alternative microfoundations may well lead to different policy implications. For example, aggregation of

Table 8
Welfare effects for socially discounting planner (COMET).

Tax policy scenario			ΔWelfare[a]			
Carbon	Capital	Labor	Benchmark	Frisch = 2.84	$\sigma = 1.1$	$\sigma = 2$
1. $\tau_{Et}^{\beta_G=.985}$	$\tau_{kt+1}^{\beta_G=.985}$	$\tau_{lt}^{\beta_G=.985}$	–	–	–	–
2. τ_{Et}^{*}	$\tau_{kt+1}^{\beta_G=.985}$	Constant	+0.96%	+1.46%	+0.93%	+0.89%
3. $\tau_{Et}^{\beta_G=.985}$	τ_{kt+1}^{*}	Constant	+1.36%	+1.97%	+1.27%	+1.21%
4. $\tau_{Et}^{\beta_G=.985}$	$\tau_{kt+1}^{\beta_G=.985}$	τ_{lt}^{*}	+2.30%	+3.72%	+2.41%	+1.76%
5. τ_{Et}^{*}	τ_{kt+1}^{*}	τ_{lt}^{*}	+3.49%	+4.94%	+3.49%	+2.85%
6. None	$\tau_{kt+1}^{\beta_G=.985}$	Constant	−4.61%	−4.17%	−5.27%	−3.24%

[a] Equivalent variation permanent percentage change in consumption.

heterogeneous rates of time preference could yield a declining representative discount rate while maintaining decentralized real interest rates (Heal and Millner, 2013). Declining total discount rates are also in line with recent empirical evidence from very long-term housing leaseholds (Giglio et al., 2014), and could be the result of uncertainty over consumption growth (see, e.g., Gollier, 2016). As another example, Gerlagh and Liski (2018) show, in a setting with quasi-hyperbolic discounting, that governments may value climate investments differently than households if delays in the climate system create a commitment value for carbon versus capital investments. In their setting, a capital income tax is required along with carbon taxes to decentralize the optimal allocation. On the other hand, many have argued for a disconnect between planner and household discounting on behavioral grounds. The policy implications in this case are likely to be similar. For example, Allcott et al. (2014) show that purchases of energy efficient durables should be subsidized if households discount the returns on these investments excessively.

Overall, the results of this paper thus demonstrate that differential discounting has far-reaching implications for both fiscal and environmental policy. While some of the results – such as the desirability of capital income subsidies – align with new research on optimal estate taxation and intergenerational insurance (Farhi and Werning, 2005, 2010), other results – such as declining labor-consumption tax wedges – represent a stark departure from conventional wisdom on optimal taxation. How governments should discount the future, and what the corresponding policy implications are, thus remains an open and essential area of research.

Acknowledgments

I thank Chris Knittel and two anonymous referees for their insightful comments and excellent suggestions. I also thank Jesse Shapiro for the inspiration for and insightful comments about this paper's research question, and Cihan Artunc, Laura Bakkensen, Mark Roberts, Kieran Walsh, and seminar participants at the Yale Energy and Environment Day, the Columbia Sustainable Development Seminar, ICMAIF 2016, and AERE 2016 for their comments.

Appendix A. Benchmark model

A.1. First-best policies: Result1

The planner's problem in the first-best setting is as follows:

$$\begin{aligned}\max & \sum_{t=0}^{\infty} \beta_G^t [U(C_t, S_t)] \\ & + \sum_{t=0}^{\infty} \beta_G^t \lambda_t [F(A_t(S_t); K_t, L_t^Y, E_t) + (1-\delta)K_t - C_t - K_{t+1}] \\ & + \sum_{t=0}^{\infty} \beta_G^t \xi_t [S_t - \widetilde{S}_t(S_0, E_{-T}, E_{-T+1}, ..., E_t)] \\ & + \sum_{t=0}^{\infty} \beta_G^t \chi_t [G_t(L_t^E) - E_t] \\ & + \sum_{t=0}^{\infty} \beta_G^t \lambda_{lt} [L_t - L_t^Y - L_t^E]\end{aligned} \tag{A.1}$$

where λ_t, ξ_t, χ_t, and λ_{lt} denote the Lagrange multipliers on the final goods resource constraint, the carbon concentrations constraint, the energy production constraint, and the labor resource constraint, respectively.

A.1.1. Decentralization: carbon taxes

In order to demonstrate that the proposed policies of Result 1 can decentralize the optimal allocation, consider first the competitive energy producer's profit-maximizing problem for a given excise emissions tax τ_{Et} at each time $t \geq 0$:

$$\begin{aligned}\max & (p_{Et} - \tau_{Et})E_t - w_t L_t^E \\ & \text{s.t.} E_t = G_t(L_t^E).\end{aligned}$$

The energy producer's optimality conditions imply that, in equilibrium,

$$p_{Et} - \tau_{Et} = \frac{w_t}{G_{Lt}} \tag{A.2}$$

Next, profit maximization by the final good producer implies that marginal products are equated to factor prices as per Eq. (10). Substituting these equilibrium prices into the energy pricing condition Eq. (A.2) yields:

$$F_{Et} - \tau_{Et} = \frac{F_{Lt}}{G_{Lt}}. \tag{A.3}$$

Finally, note that, in the social planner's problem Eq. (A.1), the FOCs with respect to labor supplied to final goods production L_t^Y and energy production L_t^E imply that the marginal cost of energy production at the optimum is given by:

$$\frac{\chi_t}{\lambda_t} = \frac{F_{Lt}}{G_{Lt}}. \tag{A.4}$$

Comparing the energy producer's optimality condition Eq. (A.2) with the planner's optimality condition for energy usage (Eq. (15)), it is thus immediately obvious that setting the carbon tax as noted in Result 1 makes the two optimality conditions coincide, as desired.

A.1.2. Decentralization: capital income taxes

Consider the representative household's Euler Eq. (8):

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H\{1 + (r_{t+1} - \delta)(1 - \tau_{kt+1})\}. \tag{A.5}$$

Noting that, in equilibrium, $r_{t+1} = F_{Kt+1}$ as per the firm's optimality condition Eq. (10), it is then straightforward to show that setting $\tau^*_{kt+1} = \left(\frac{\beta_H - \beta_G}{\beta_H}\right)\left(\frac{F_{Kt+1}+1-\delta}{F_{Kt+1}-\delta}\right)$ makes the household's intertemporal optimality condition coincide with that of the social planner (Eq. (16)).

A.2. Constrained-optimal climate policy: Proposition1

Letting Φ_t denote the Lagrange multiplier on the no-subsidy constraint Eq. (21) added to the planner's problem Eq.(A.1), the planner's optimality condition for energy usage E_t yields the following revised version of Eq. (15) defining the optimal carbon tax τ^*_{Et}:

$$\underbrace{\underbrace{\underbrace{F_{Et}}_{\substack{\text{Marginal}\\ \text{product of}\\ \text{energy}}} - \underbrace{\sum_{j=0}^{\infty}\beta_G^j\frac{\partial S_{t+j}}{\partial E_t}\left[\frac{\xi_{t+j}}{\lambda_t}\right]}_{\substack{\text{SDV of marginal}\\ \text{damages from}\\ \text{carbon emissions}}} + \underbrace{\frac{\Phi_{t-1}}{\lambda_t\beta_G}\frac{\partial F_{Kt}}{\partial E_t}}_{\substack{\text{Marginal energy use}\\ \text{impact on marginal}\\ \text{product of capital}}}}_{=-\tau^*_{Et}} = \underbrace{\frac{\chi_t}{\lambda_t}}_{\substack{\text{MC of}\\ \text{energy production}\\ (\text{in} C_t \text{ units})}}. \tag{A.6}$$

Optimal energy usage thus balances its benefits in final goods production (F_{Et}) minus the externality cost of the associated climate change $\left(\sum_{j=0}^{\infty}\beta_G^j\frac{\partial S_{t+j}}{\partial E_t}\left[\frac{\xi_{t+j}}{\lambda_t}\right]\right)$ and the social cost of the effect of energy input usage on the marginal product of capital and thus investment incentives $\left(\frac{\Phi_{t-1}}{\lambda_t\beta_G}\frac{\partial F_{Kt}}{\partial E_t}\right)$, against the marginal costs of producing energy $\left(\frac{\chi_t}{\lambda_t}\right)$. Taking the optimality conditions with respect to labor in each sector, L_t^Y and L_t^E, confirms that $\frac{\chi_t}{\lambda_t}$ continues to equal the energy producer's production cost as in Eq.(A.4). Next, the externality cost of carbon is defined by the FOC with respect to the stock S_t:

$$\underbrace{-\xi_t}_{\substack{\text{Shadow}\\ \text{value of carbon}\\ \text{concentrations}}} = \underbrace{\lambda_t\frac{\partial Y_{t+j}}{\partial S_{t+j}} + U_{St}}_{\substack{\text{Output and}\\ \text{disutility}\\ \text{damages}}} + \underbrace{\frac{\Phi_{t-1}}{\beta_G}\frac{\partial F_{Kt}}{\partial S_t}}_{\substack{\text{Marginal damages}\\ \text{to marginal product}\\ \text{of capital}}}. \tag{A.7}$$

Substituting Eq. (A.7) into Eq. (A.6), iterating forward, and comparing the resulting expression with the energy producer's profit-maximizing conditions in competitive equilibrium reveals that the constrained-optimal tax is implicitly defined by the desired expression from Proposition 1. □.

Appendix B. Distortionary fiscal setting

B.1. Proof of Proposition2

Claim 1. *"If preferences satisfy Eq.(A), then optimal fiscal policy creates an effective savings subsidy for all $t > 1$ (i.e., a negative wedge between the intertemporal marginal rate of transformation and substitution)."*

Proof. Taking and rearranging the first-order conditions of the social planner's problem readily leads to the intertemporal optimality condition Eq. (29):

$$\frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^t U_{ct} + \phi W_{ct}\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} U_{ct+1} + \phi W_{ct+1}\right]}\frac{1}{\beta_H} = [(1 - \delta) + F_{Kt+1}]. \tag{B.1}$$

If preferences are of the form Eq.(A), it is straightforward to show that $W_{ct} = 0$ for all $t > 0$, so that the optimality condition Eq. (B.1) becomes:

$$\underbrace{\frac{U_{ct}}{U_{ct+1}\beta_H}}_{\equiv MRS_{t,t+1}} = \underbrace{[(1 - \delta) + F_{Kt+1}]}_{\equiv MRT_{t,t+1}}\left(\frac{\beta_G}{\beta_H}\right). \tag{B.2}$$

Given the relevant definitions of the intertemporal marginal rate of transformation ($MRT_{t,t+1}$) and substitution ($MRS_{t,t+1}$), Eq. (B.2) readily implies the desired result (given that $\beta_H < \beta_G$ by assumption):

$$MRT_{t,t+1} - MRS_{t,t+1} = [(1 - \delta) + F_{Kt+1}]\left(\frac{\beta_H - \beta_G}{\beta_H}\right) < \quad 0.$$

Claim 2. *"If preferences satisfy Eqs.(A) and(A'), then the optimal fiscal policy creates an effective labor-consumption tax wedge that is decreasing over time for all $t > 1$."*

Proof. If preferences satisfy Eq.(A') in addition to Eq.(A), then $W_{lt} = U_{lt}v_2$. Substituting into the planner's intratemporal optimality condition Eq. (28) and rearranging yields:

$$\frac{-U_{lt}}{U_{ct}} = F_{Lt}\left[1 + \nu_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t\right]^{-1}. \tag{B.3}$$

Next, consider the household's intratemporal optimality condition Eq. (25). In equilibrium, $w_t = F_{Lt}$. Further invoking the definition of an effective labor-consumption tax wedge τ_{lct} via $(1-\tau_{lct}) \equiv \frac{(1-\tau_{lt})}{(1+\tau_{ct})}$, the household's optimality condition becomes:

$$\frac{-U_{lt}}{U_{ct}} = F_{Lt}(1-\tau_{lct}). \tag{B.4}$$

Comparison of Eqs. (B.3) and (B.4) immediately reveals that the effective labor-consumption tax wedge decentralizing the optimal allocation satisfies (for $t > 0$):

$$\tau^*_{lct} = 1 - \left[1 + \nu_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t\right]^{-1}.$$

The fact that τ^*_{lct} is decreasing over time then follows straightforwardly given that (i) $\nu_2 > 1$ by assumption, (ii) the Lagrange multiplier on Eq.(IMP) is $\phi > 0$, and (iii) $\beta_H < \beta_G$. □.

B.2. Proof of Corollary1

Corollary 1 describes two specific decentralizations of the optimal allocation: One employing capital, labor, and emissions taxes ("KLE decentralization"), and the other employing capital, consumption, and emissions taxes ("KCE decentralization ").

B.2.1. KLE decentralization

Without consumption taxes, the household's optimality condition Eqs. (25)–(26) become:

$$\frac{-U_{lt}}{U_{ct}} = w_t(1-\tau_{lt}) \tag{B.5}$$

$$\frac{U_{ct}}{U_{ct+1}\beta_H} = \{1 + (r_{t+1} - \delta)(1-\tau_{kt+1})\}. \tag{B.6}$$

First, invoking that $r_{t+1} = F_{Kt+1}$ in equilibrium and comparing Eq. (B.6) with the planner's intertemporal optimality condition Eq. (B.2) reveals that the capital income tax required to decentralize the optimal allocation (for $t > 1$) is defined by:

$$\tau^*_{kt+1} = \left(\frac{\beta_H - \beta_G}{\beta_H}\right)\frac{(F_{Kt+1} - \delta + 1)}{(F_{Kt+1} - \delta)}. \tag{B.7}$$

With standard discounting ($\beta_H = \beta_G$), the optimal capital income tax (for $t > 1$) is thus zero. With social discounting ($\beta_H < \beta_G$), the optimal capital income tax is thus negative, i.e., a subsidy.[10]

Next, invoking that $w_t = F_{Lt}$ in equilibrium and comparing Eq. (B.5) with the planner's intratemporal optimality condition Eq. (B.3) reveals that the labor income tax required to decentralize the optimal allocation (for $t > 0$) is defined by:

$$\tau^*_{lt} = 1 - \left[1 + \nu_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t\right]^{-1}.$$

With standard discounting ($\beta_H = \beta_G$), the optimal labor income tax is thus constant over time and positive (given that (i) $\nu_2 > 1$ by assumption, and that (ii) the Lagrange multiplier on the (IMP) is $\phi > 0$). With social discounting ($\beta_H < \beta_G$), the optimal labor income tax is thus decreasing over time but remains weakly positive (as τ^*_{lt} asymptotes to zero as $t \to \infty$). □.

B.2.2. KCE decentralization

Without labor taxes, the household's optimality condition Eqs. (25)–(26) become:

$$\frac{-U_{lt}}{U_{ct}} = \frac{w_t}{1+\tau_{ct}} \tag{B.8}$$

$$\frac{U_{ct}}{U_{ct+1}\beta_H} = \{1 + (r_{t+1} - \delta)(1-\tau_{Kt+1})\}\frac{(1+\tau_{ct})}{(1+\tau_{ct+1})}. \tag{B.9}$$

First, invoking that $w_t = F_{Lt}$ in equilibrium and comparing Eq. (B.8) with the planner's intratemporal optimality condition Eq. (B.3) reveals that the consumption tax required to decentralize the optimal allocation (for $t > 0$) is defined by:

$$\tau^*_{ct} = \nu_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t. \tag{B.10}$$

With standard discounting ($\beta_H = \beta_G$), the optimal consumption tax is thus constant over time and positive. With social discounting ($\beta_H < \beta_G$), the optimal consumption tax is decreasing over time but remains weakly positive (as τ^*_{ct} asymptotes zero as $t \to \infty$).

[10] Of course this assumes that the marginal product of capital net of depreciation is positive.

Next, given the optimal consumption tax Eq. (B.10) and with the fact that $r_{t+1} = F_{Kt+1}$ in equilibrium, the household's intertemporal optimality condition Eq. (B.9) in this setting becomes:

$$\frac{U_{ct}}{U_{ct+1}\beta_H} = \{1 + (F_{Kt+1} - \delta)(1 - \tau_{kt+1})\}\left[\frac{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^t}{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}}\right]. \tag{B.11}$$

Comparing Eq. (B.11) with the planner's intertemporal optimality condition Eq. (B.2) then yields an implicit expression for the optimal capital income tax in this setting:

$$\left(\frac{F_{Kt+1} - \delta + 1}{F_{Kt+1} - \delta}\right)\left[1 - \left(\frac{\beta_G}{\beta_H}\right)\left[\frac{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}}{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t}}\right]\right] = \tau^*_{kt+1}. \tag{B.12}$$

With standard discounting ($\beta_H = \beta_G$), the square bracketed term in Eq. (B.12) reduces to zero, indicating that the optimal capital income tax (for $t > 1$) is zero. With social discounting ($B_H < \beta_G$), the optimal capital income tax is negative as:

$$\begin{aligned} 1 \;&>\; \frac{\frac{\beta_H}{\beta_G} + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}}{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}} \\ \Rightarrow \left(\frac{\beta_G}{\beta_H}\right) &> \frac{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t}}{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}} \\ \Rightarrow \left(\frac{\beta_G}{\beta_H}\right)\left[\frac{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t+1}}{1 + v_2\phi\left(\frac{\beta_H}{\beta_G}\right)^{t}}\right] &> \quad 1. \end{aligned} \tag{B.13}$$

The last inequality Eq. (B.13) implies that the square bracketed term in Eq. (B.12) will be negative, and, consequently, that the optimal capital income tax is negative as well. □.

Appendix C. GHKT-B recalibration

C.1. Energy sector

The benchmark numerical framework uses a simplified energy sector representation compared to GHKT, who allow for multiple energy inputs and non-renewable resource dynamics. Whereas GHKT calibrate separate labor productivity parameters for coal and oil production, here we need a representative figure for both. In the benchmark scenario in the base decade 2010, the GHKT model yields a share of oil in total fossil energy consumption (all in tons of carbon-equivalent) of (33.6/55.2) = 61%. Similarly, the IEA reports that oil and gas accounted for 55% of global CO_2 emissions in 2012 (International Energy Agency, 2012). GHKT compute model base year prices of oil/gas and coal of *$606.5/mtC* and *$103.5/mtC*, respectively. The weighted average price is thus *$405.3/mtC*. I thus calculate base year energy sector productivity as:

$$A_{E0}(\$/GtC) = \frac{w_0}{p_{E0}} \sim \frac{(1 - \alpha - v)Y_0}{\$405.3{\cdot}10^9} = \quad 1140.$$

C.2. CES production

The GHKT framework focuses on Cobb-Douglas aggregate production, where the elasticity of substitution between energy and capital inputs is equal to unity. In order to illustrate the importance of energy-capital complementarity for the constrained-optimal carbon tax, I introduce an alternative nested CES production function:

$$Y_t = (1 - D(S_t)){\cdot}A_t\left[\varkappa[K_t^{\alpha}L_t^{1-\alpha}]^{\frac{\varepsilon-1}{\varepsilon}} + (1 - \varkappa)[E_t]^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

This specification is motivated by Van der Werf (2008), who, in reviewing the literature as well as cross-country data, concludes that a (KL)E nesting specification generally provides the best fit for climate-economy models. In line with several other studies, I consider $\varepsilon_{KL,E} = 0.4$ as a benchmark case. Given base period output shares, the distribution parameter is set to $\varkappa = 0.96$. Base period labor supply is normalized to $L_0 = 1$, the initial net damage term can be calculated as $(1 - D(S_0)) = 0.9948$, and base period energy inputs are given from the data and the GHKT calibration at $E_0 = 34.9522$ *GtC* per decade. Note that I retain the initial capital stock measure K_0 from the benchmark model (which is calibrated such that the marginal product of capital in the Cobb-Douglas production framework matches the decadal net return to capital for an annual return equivalent of 5% plus 100% decadal depreciation). Finally, initial TFP A_0 can then be calibrated as:

$$A_0 = \frac{Y_0}{(1 - D(S_0))\left[\varkappa[K_0^{\alpha}L_0^{1-\alpha}]^{\frac{\varepsilon-1}{\varepsilon}} + (1 - \varkappa)[E_0]^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}}.$$

C.3. Optimization

Given the high discount factor, I extend the model's direct optimization period to 50 periods (500 years), after which savings rates are locked in and iterated forward for 1000 years for the climate to achieve a steady state, after which a balanced growth path is imposed to compute the continuation value of the allocations over an infinite time horizon. For further details on the computation, see GHKT (2014) and Barrage (2014).

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpubeco.2018.02.012.

References

Acemoglu, D., Golosov, M., Tsyvinski, A., 2011. Political economy of Ramsey taxation. J. Public Econ. 95 (7), 467–475.

Allcott, H., Mullainathan, S., Taubinsky, D., 2014. Energy policy with externalities and internalities. J. Public Econ. 112, 72–88.

Arrow, K.J., Cropper, M.L., Gollier, C., Groom, B., Heal, G.M., Newell, R.G., Nordhaus, W.D., Pindyck, R.S., Pizer, W.A., Portney, P.R., Sterner, T., Tol, R.S.J., Weitzman, M.L., 2012. How should benefits and costs be discounted in an intergenerational context? In: Resources for the Future Discussion Paper 12–42.

Arrow, K.J., Cline, W.R., Maler, K.G., Munasinghe, M., Squitieri, R., Stiglitz, J.E., 1996. Intertemporal Equity, Discounting, and Economic Efficiency, intergovernmental Panel on Climate Change. In: Intergovernmental Panel on Climate Change (IPCC) Second Assessment Report, Working Group II, 1996.

Asheim, G.B., Nesje, F., 2016. Destructive intergenerational altruism. J. Assoc. Environ. Resour. Econ. 3 (4), 957–984.

Atkeson, A., Chari, V.V., Kehoe, P.J., 1999. Taxing capital income: a bad idea. Fed. Reserve Bank Minneap. Q. Rev. 23 (3), 3–17.

Barrage, L., 2014. Sensitivity analysis for Golosov, Hassler, Krusell, and Tsyvinski (2013): optimal taxes on fossil fuel in general equilibrium, Supplementary Material. Econometrica 82 (1), 41–88 [Supplemental Material to: Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) "Optimal Taxes on Fossil Fuel in General Equilibrium," Econometrica, 82(1)].

Barrage, L., 2016. Optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy. In: Brown University Working Paper.

Baumol, W.J., 1968. On the social rate of discount. Am. Econ. Rev. 58 (4), 788–802.

Belfiori, M.E., 2017. Carbon pricing, carbon sequestration, and social discounting. Eur. Econ. Rev. 96, 1–17.

Benhabib, J., Rustichini, A., 1997. Optimal taxes without commitment. J. Econ. Theory 77, 231–259.

Bernheim, B.D., 1989. Intergenerational altruism, dynastic equilibria and social welfare. Rev. Econ. Stud. 56 (1), 119–128.

Bovenberg, A.L., Goulder, L.H., 1996. Optimal environmental taxation in the presence of other taxes: general-equilibrium analyses. Am. Econ. Rev. 86 (4), 985–1000.

Caplin, A., Leahy, J., 2004. The social discount rate. J. Polit. Econ. 112, 1257–1268.

Chari, V.V., Kehoe, P.J., 1999. Optimal fiscal and monetary policy. Handb. Macroecon. 1, 1671–1745.

Chetty, R., Guren, A., Manoli, D., Weber, A., 2011. Micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. Am. Econ. Rev. Pap. Proc. 101 (3), 471–475.

Cline, W.R., 1992. The Economics of Global Warming Washington. Institute for International Economics, Washington, DC.

Dasgupta, P., 2008. Discounting climate change. J. Risk Uncertain. 37 (2-3), 141–169.

Farhi, E., Werning, I., 2005. Inequality, social discounting, and estate taxation. In: NBER WP 11408.

Farhi, E., Werning, I., 2007. Inequality and social discounting. J. Polit. Econ. 115 (3), 365–402.

Farhi, E., Werning, I., 2010. Progressive estate taxation. Q. J. Econ. 125 (2), 635–673.

Gerlagh, R., Liski, M., 2018. Consistent climate policies. J. Eur. Econ. Assoc. 16 (1), 1–44 Jvx010.

Giglio, S., Maggiori, M., Stroebel, J., 2014. Very long-run discount rates. Q. J. Econ. 130 (1), 1–53.

Gollier, C., 2016. Evaluation of long-dated assets: the role of parameter uncertainty. J. Monet. Econ. 84, 66–83.

Gollier, C., Weitzman, M.L., 2010. How should the distant future be discounted when discount rates are uncertain? Econ. Lett. 107 (3), 350–353.

Gollier, C., Zeckhauser, R., 2005. Aggregation of heterogeneous time preferences. J. Polit. Econ. 113 (4), 878–896.

Golosov, M., Hassler, J., Krusell, P., Tsyvinski, A., 2014. Optimal taxes on fossil fuel in general equilibrium. Econometrica 82 (1), 41–88.

Goulder, L.H., 1995. Environmental taxation and the double dividend: a reader's guide. Int. Taw Public Financ. 2, 157–183.

Goulder, L.H., Williams, R. III, 2012. The choice of discount rate for climate change policy evaluation. Clim. Chang. Econ. 3 (4) 1250024-1-1250024-18.

Heal, G., Millner, A., 2013. Discounting under disagreement. In: NBER WP No. w18999.

Interagency Working Group on Social Cost of Carbon, 2010. Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866. United States Government.

International Energy Agency, 2012. 2012 CO_2 Emissions Overview. available at [last accessed June 2016]: https://www.iea.org/media/statistics/topics/emissions/CO2_Emissions_Overview.pdf.

Iverson, T., Karp, L., 2017. Carbon taxes and commitment with non-constant time preference. In: Working Paper.

Jones, L.E., Manuelli, R.E., Rossi, P.E., 1993. Optimal taxation in models of endogenous growth. J. Polit. Econ. 101 (3), 485–517.

Judd, K.L., 1985. Redistributive taxation in a simple perfect foresight model. J. Public Econ. 28, 59–83.

Judd, K.L., 1999. Optimal taxation and spending in general competitive growth models. J. Public Econ. 71, 1–26.

Kaplow, L., Moyer, E., Weisbach, D.A., 2010. The social evaluation of intergenerational policies and its application to integrated assessment models of climate change. BE J. Econ. Anal. Policy 10 (2).

Klein, P., Ríos-Rull, J.V., 2003. Time-consistent optimal fiscal policy. Int. Econ. Rev. 44 (4), 1217–1245.

King, R.G., Plosser, C.I., Rebelo, S.T., 2001. Production, growth and business cycles: technical appendix. Comput. Econ. 20 (1-2), 87–116.

Lemoine, D.M., Traeger, C., 2014. Watch your step: optimal policy in a tipping climate. Am. Econ. J.: Econ. Policy 6 (1), 137–166.

Lind, R.C. (Ed.), 1982. Discounting for time and risk in energy policy. Resources for the Future, Washington, DC.

Manne, A.S., 1995. The rate of time preference — implications for the greenhouse debate. Energy Policy 23 (4/5), 391–394.

Newell, R.G., Pizer, W.A., 2003. Discounting the distant future: how much do uncertain rates increase valuations? J. Environ. Econ. Manag. 46 (1), 52–71.

Nordhaus, W.D., 2007. A review of the "Stern review on the economics of climate change". J. Econ. Lit. XLV, 686–702.

Nordhaus, W.D., 2008. A Question of Balance. Yale University Press, New Haven, CT.

Nordhaus, W.D., 2010a. DICE-2010 Model. available at http://nordhaus.econ.yale.edu/RICEmodels.htm.

Nordhaus, W.D., 2010b. Economic aspects of global warming in a post-Copenhagen environment. Proc. Natl. Acad. Sci. 107 (26), 11721–11726.

Ramsey, F.P., 1928. A mathematical theory of saving. Econ. J. 38 (152), 543–559.

Ryan, S.P., 2012. The costs of environmental regulation in a concentrated industry. Econometrica 80 (3), 1019–1061.

Schmitt, A., 2014. Optimal carbon and income taxation. In: Working Paper.

Sen, A.K., 1982. Approaches to the choice of discount rates for social benefit-cost analysis. In: Lind, R.C. (Ed.), Resources for the Future, Washington, DC.

Sleet, C., Yeltekin, Ş., 2006. Credibility and endogenous societal discounting. Rev. Econ. Dyn. 9 (3), 410–437.

Stern, N.H., 2006. Stern Review: The Economics of Climate Change. vol. 30. HM Treasury, London, pp. 2006.

Sterner, T., Persson, U.M., 2008. An even sterner review: introducing relative prices into the discounting debate. Rev. Environ. Econ. Policy 2 (1), 61–76.

Traeger, C., 2011. Sustainability, limited substitutability and non-constant social discount rates. J. Environ. Econ. Manag. 62, 215–228.

Van der Werf, E., 2008. Production functions for climate policy modeling: an empirical analysis. Energy Econ. 30 (6), 2964–2979.

von Below, D., 2012. Optimal Carbon Taxes with Social and Private Discounting. Working Paper, SURED Conference, Ascona, Switzerland.